LAV$_{ELI}$
Attorney Docket
No. GUPL

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that: my residence, post office address and citizenship are as stated below next t my name; that I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joir inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the inventio entitled: A VARIANT OF LAV VIRUSES

the specification of which is ☒ attached and/or ☒ was filed on ..13 April 1987.......... as Application Serial No ......038.332.......... and was amended on (if applicable) ....

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

COUNTRY	APPLICATION NUMBER	DATE OF FILING	PRIORITY CLAIMED UNDER 35 U.S.C. 119
European Patent Convention	86401380.0	23 June 1986	☒ YES ☐ NO
			☐ YES ☐ NO

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

APPLICATION NUMBER	DATE OF FILING	STATUS (Patented, Pending, Abandoned)

I hereby appoint the following attorneys to prosecute this application and transact all business in the Patent and Trademark Office connected therewith: Finnegan, Henderson, Farabow, Garrett and Dunner, Reg. No. 22,540; Douglas B. Henderson, Reg. No. 20,291; Ford. F. Farabow, Jr., Reg. No. 20,630; Arthur S. Garrett, Reg. No. 20,338; Donald R. Dunner, Reg. No. 19,073; Brian G. Brunsvold, Reg. No. 22,593; Tipton D. Jennings, IV, Reg. No. 20,645; Jerry D. Voight, Reg. No. 23,020; Laurence R. Hefter, Reg. No. 20,827; Kenneth E. Payne, Reg. No. 23,098; Herbert H. Mintz, Reg. No. 26,691; C. Larry O'Rourke, Reg. No. 26,014; Albert J. Santorelli, Reg. No. 22,610; Michael C. Elmer, Reg. No. 25,857; Richard H. Smith, Reg. No. 20,609; Stephen L. Peterson, Reg. No. 26,325; John M. Romary, Reg. No. 26,331; Bruce C. Zotter, Reg. No. 27,680; Dennis P. O'Reilley, Reg. No. 27,932; Allen M. Sokal, Reg. No. 26,695; Robert D. Bajefsky, Reg. No. 25,387; Richard L. Stroup, Reg. No. 28,478; David W. Hill, Reg. No. 28,220; Thomas L. Irving, Reg. No. 28,619; Charles E. Lipsey, Reg. No. 28,165; Thomas W. Winland, Reg. No. 27,605; and .... Please address all correspondence to FINNEGAN, HENDERSON, FARABOW, GARRETT AND DUNNER, 1775 K Street, N.W., Washington, D.C. 20006, Telephone No. (202) 293-6850.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Listing of Inventors Con[redacted]on Page 2 hereof. ☐ Yes ☐ No

FULL NAME OF SOLE OR [redacted]NTOR	INVENTOR'S SIGNATURE	DATE
Alizon, Mar[redacted]		4.5.87

RESIDENCE	CITIZENSHIP
71, rue de Cardinal Lemoine, Paris 75005 France	France

POST OFFICE ADDRESS
same

FULL NAME OF SECOND JOINT INVENTOR, IF ANY	INVENTOR'S SIGNATURE	DATE
Sonigo, Pierre		29/4/87

RESIDENCE	CITIZENSHIP
23, rue Gutenberg, Paris 75015 France	France

POST OFFICE ADDRESS
same

FULL NAME OF THIRD JOINT INVENTOR, IF ANY	INVENTOR'S SIGNATURE	DATE
Wain-Hobson, Simon	Simon Wain-Hobson	5/5/87

RESIDENCE	CITIZENSHIP
3, rue Jean de la Fontaine, Montigny Les Bretonneux 78180 France	Great Britain

POST OFFICE ADDRESS
same

FULL NAME OF SOLE OR FIRST INVENTOR
Montagnier, Luc

INVENTOR'S SIGNATURE

DATE
27/3/87

RESIDENCE
21, rue de Malabry, Le Plessis Robinson 92350 France

CITIZENSHIP
France

POST OFFICE ADDRESS
same

FULL NAME OF SECOND JOINT INVENTOR, IF ANY
INVENTOR'S SIGNATURE
DATE
RESIDENCE
CITIZENSHIP
POST OFFICE ADDRESS

FULL NAME OF THIRD JOINT INVENTOR, IF ANY
INVENTOR'S SIGNATURE
DATE
RESIDENCE
CITIZENSHIP
POST OFFICE ADDRESS

FULL NAME OF SOLE OR FIRST INVENTOR
INVENTOR'S SIGNATURE
DATE
RESIDENCE
CITIZENSHIP
POST OFFICE ADDRESS

FULL NAME OF SECOND JOINT INVENTOR, IF ANY
INVENTOR'S SIGNATURE
DATE
RESIDENCE
CITIZENSHIP
POST OFFICE ADDRESS

FULL NAME OF THIRD JOINT INVENTOR, IF ANY
INVENTOR'S SIGNATURE
DATE
RESIDENCE
CITIZENSHIP
POST OFFICE ADDRESS

FULL NAME OF SECOND JOINT INVENTOR, IF ANY
INVENTOR'S SIGNATURE
DATE
RESIDENCE
CITIZENSHIP
POST OFFICE ADDRESS

FULL NAME OF THIRD JOINT INVENTOR, IF ANY
INVENTOR'S SIGNATURE
DATE
RESIDENCE
CITIZENSHIP
POST OFFICE ADDRESS

PATENT
Attorney Docket No. 2356.0010-04

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)
Marc ALIZON et al.	)
Application No.: Unassigned (Cont. of U.S.S.N. 08/423,477 (4/19/95))	) Group Art Unit: Unassigned
Filed: January 23, 2001	) Examiner: Unknown
For: VARIANT OF LAV VIRUSES	)

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

SUBMISSION OF FORMAL DRAWINGS

Subject to the approval of the Examiner, applicants submit the attached 34 sheets of Formal Drawings (Figs. 1A, 1B, 2, 3A-1, 3A-2, 3B-1, 3B-2, 3C-1, 3C-2, 3D-1, 3D-2, 3E-1, 3E-2, 3F-1, 3F-2, 4A, 4B, 5, 6A-1, 6A-2, 6A-3, 6B-1, 6B-2, 6B-3, 6B-4, 7A, 7B, 7C, 7D, 7E, 7F, 7G, 7H, and 7I). If the Formal Drawings for any reason are not in full compliance with the pertinent statutes and regulations, please so advise the undersigned.

Please grant any extensions of time required to enter this response and charge any additional required fees to our deposit account 06-0916.

Respectfully submitted,

FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, L.L.P.

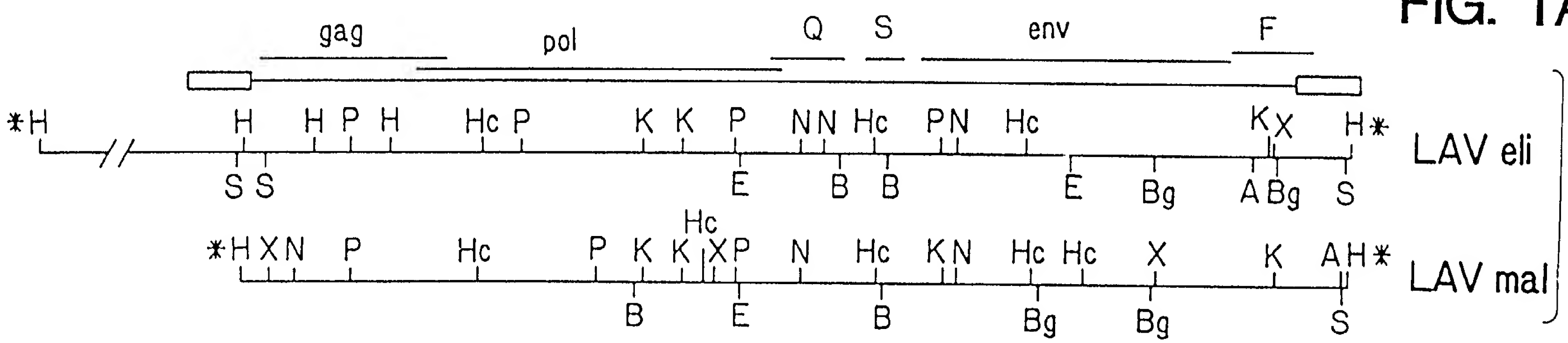
By: ____________________

Salvatore J. Arrigo
Reg. No. 46,063
Tel: (202) 408-4000
Fax: (202) 408-4400
Email: arrigos@finnegan.com

Date: January 23, 2001

LAW OFFICES
FINNEGAN, HENDERSON, FARABOW, GARRETT, & DUNNER, L.L.P.
1300 I STREET, N. W.
WASHINGTON, DC 20005
202-408-4000

FIG. 1A


gag
pol
Q
S
env
F
*H
H H P H Hc P K K P N N Hc P N Hc K X H*
S S E B B E Bg A Bg S
LAV eli
H X N P Hc P K K Hc X P N Hc K N Hc Hc X K A H
B E B Bg Bg S
LAV mal

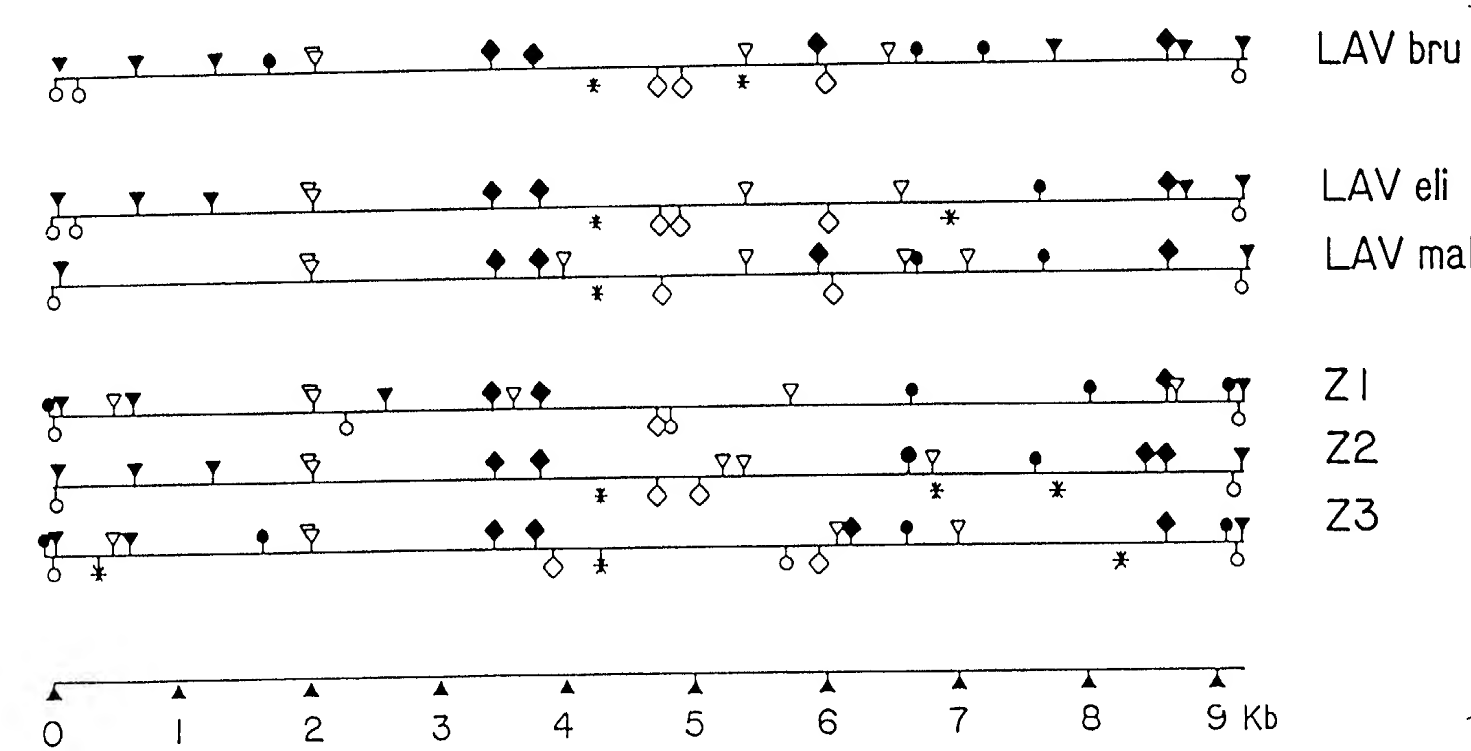
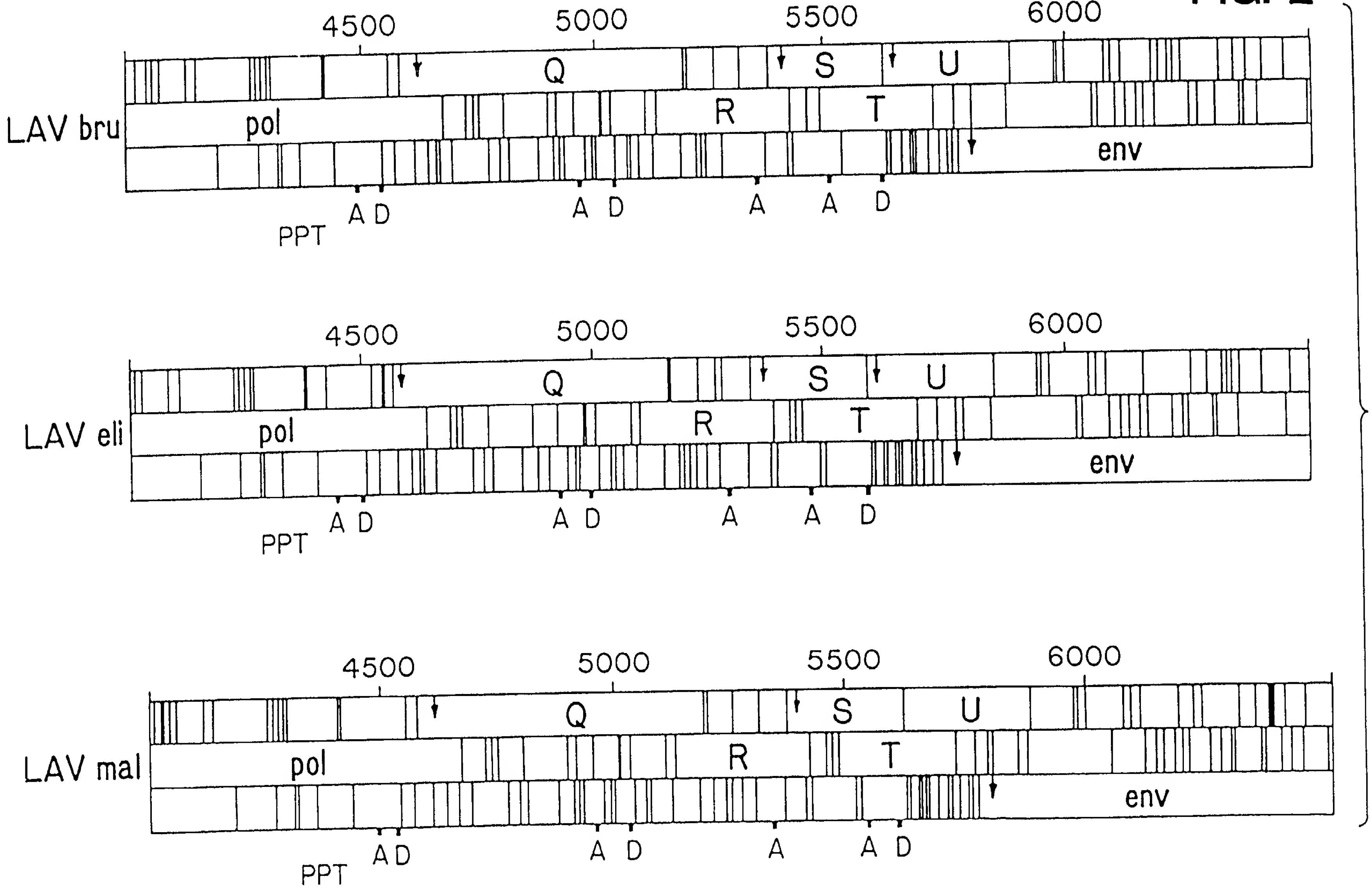


FIG. 1B

FIG. 2



4500
5000
5500
6000
LAV bru
LAV eli
LAV mal
pol
Q
S
U
R
T
env
A D
A D
A
A
D
PPT

GAG

```
                 10         20         30         40         50         60         70         80
LAV BRU  MGARASVLSG GELDRWEKIR LRPGGKKKYK LKHIVWASRE LERFAVNPGL LETSEGCRQI LGQLQPSLQT GSEELRSLYN
ARV 2                   K                    L            L        C    Q  ME   ST  K      IK
LAV MAL              K  A               R               Y L             K   I    AI     T
LAV ELI              K  K               R

                                                                        p25
                                                                         ↓
                 90        100        110        120        130        140        150        160
LAV BRU  TVATLYCVHQ RIEIKDTKEA LDKIEEEQNK SKKKAQQAAA ------DTGH SSQVSQNYPI VQNIQGQMVH QAISPRTLNA
ARV 2                  DV       E                    ----AAG  N               L
LAV MAL                DV            I     RQ T      AQQAAAA KN   S           A     I
LAV ELI           K G  DV       E M                  -------  N N             L

                170        180        190        200        210        220        230        240
LAV BRU  WVKVVEEKAF SPEVIPMFSA LSCGATPQDL NTMLNTVGGH QAAMQMLKET INEEAAEWDK VHPVHAGPIA PGQMREPRGS
ARV 2                                      M   I             D        D             P
LAV MAL      I                                                             L
LAV ELI      I

                250        260        270        280        290        300        310        320
LAV BRU  DIAGTTSTLQ EQIGWMTNNP PIPVGEIYKR WIILGLNKIV RMYSPTSILD IRQGPKEPFR DYVDRFYKTL RAEQASQEVK
ARV 2                                                      V                     F         T D
LAV MAL                   S        D        V             V                                 D
LAV ELI               A   S
```

FIG. 3A-1

```
                                                                                      p13
                                                                                      ↓
         330        340        350        360        370        380        390        400
LAV BRU  NWMTETLLVQ NANPDCKTIL KALGPAATLE EMMTACQGVG GPGHKARVLA EAMSQVTNS- ATIMMQRGNF RNQRKIVKCF
                                                                       P-  N              T
ARV 2                                                           A     T   A        KG - RI
LAV MAL                             G                  S        A     V T A        KGP   I
LAV ELI                             Q                  S

         410        420        430        440        450        460        470        480
LAV BRU  NCGKEGHIAR NCRAPRKKGC WKCGKEGHQM KDCTERQANF LGKIWPSYKG RPGNFLQSRP EPTAPPFLQS RPEPTAPPEE
                  K            R  R                                             ----  ---------
ARV 2                                                        H                   ----  ---------A
LAV MAL          L             R       L               R     H                   ----  ---------A
LAV ELI           K

         490        500        510
LAV BRU  SFRSGVETTT PSQKQEPIDK ELYPLTSLRS LFGNDPSSQ
           F E K
ARV 2    GF E IK-     QK         A  K         QL
LAV MAL  GF E I -     QK            K          L
LAV ELI
```

FIG. 3A-2

Central Region: Q

```
                  10         20         30         40         50         60         70         80
LAV BRU   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSGKARGWFY RHHYESPHPR ISSEVHIPLG DARLVITTYW GLHTGERDWH
ARV 2                                      I K  K           T    V            K               E
LAV MAL                          H           K  KN          R  K V               VR      Q   K
LAV ELI                      K               K  NR             K          E     K           E

                  90        100        110        120        130        140        150        160
LAV BRU   LGQGVSIEWR KKRYSTQVDP ELADQLIHLY YFDCFSDSAI RKALLGHIVS PRCEYQAGHN KVGSLQYLAL AALITPKKIK
ARV 2          A       K        G        H       E    KN I  YR                                 T
LAV MAL     H        Q      L   D                E     Q I          D                  T   A   TR
LAV ELI               R         G       M        E       I  D                          T   A  Q

                 170        180        190
LAV BRU   PPLPSVTKLT EDRWNKPQKT KGHRGSHTMN GH
ARV 2           K
LAV MAL         R            Q
LAV ELI         R            Q  R
```

FIG. 3B-1

```
               R

                  10         20         30         40         50         60         70         80
LAV BRU   MEQAPEDQGP QREPHNEWTL ELLEELKNEA VRHFPRIWLH GLGQHIYETY GDTWAGVEAI IRILQQLLFI HFRIGCRHSR
ARV 2                    Y             R         P    S   Y                                  Q
LAV MAL        A                       Q              S              E        S              Q
LAV ELI        A         Y   A         S              S              V                       Q

                  90
LAV BRU   IGVTQQRRAR -NGASRS
ARV 2       II       R
LAV MAL     I R      -  S
LAV ELI     IIR      -  S

               S (tat)
                  10         20         30         40         50         60         70
LAV BRU   MEPVDPRLEP WKHPGSQPKT ACTTCYCKKC CFHCQVCFTT KALGISYGRK KRRQRRRPPQ GSQTHQVSLS KQ
ARV 2           N            R    NN           YA   R  G                A   D     A
LAV MAL    D    N     N      R  P NK        Y   M  I   G                     N A  DP P E
LAV ELI    D    N     N      R  P NK H      Y  P   LN  G               G     G A   PIP
```

FIG. 3B-2

POL

```
                10         20         30         40         50         60         70         80
LAV BRU  FFREDLAFLQ GKAREFSSEQ TRANSPTFSS EQTRANSPTR RELQVWGRDN NSLSEAGADR QGTVSFNFPQ ITLWQRPLVT
ARV 2                                 --- ----------        GE
LAV MAL      N   P        P           --- ---------S    R   G - KT   T  E   I    S          V
LAV ELI      N   P     G L PK         --- ---------S    R     -  P  KT  E                    A

                90        100        110        120        130        140        150        160
LAV BRU  IKIGGQLKEA LLDTGADDTV LEEMSLPGRW KPKMIGGIGG FIKVRQYDQI LIEICGHKAI GTVLVGPTPV NIIGRNLLTQ
ARV 2     R                       N   K                       PV
LAV MAL  VRV                     IN   K                                   K    I             M
LAV ELI                           N   K                       P     Q

               170        180        190        200        210        220        230        240
LAV BRU  IGCTLNFPIS PIETVPVKLK PGMDGPKVKQ WPLTEEKIKA LVEICTEMEK EGKISKIGPE NPYNTPVFAI KKKDSTKWRK
ARV 2
LAV MAL                             R                 T    KD      L
LAV ELI                                               T     D       R           I

               250        260        270        280        290        300        310        320
LAV BRU  LVDFRELNKR TQDFWEVQLG IPHPAGLKKK KSVTVLDVGD AYFSVPLDED FRKYTAFTIP SINNETPGIR YQYNVLPQGW
ARV 2                                                        K
LAV MAL    N
LAV ELI                                                                S
```

FIG. 3C-1

```
                330        340        350        360        370        380        390        400
LAV BRU  KGSPAIFQSS MTKILEPFRK QNPDIVIYQY MDDLYVGSDL EIGQHRTKIE ELRQHLLRWG LTTPDKKHQK EPPFLWMGYE
ARV 2                                                                       F
LAV MAL                      T K  E                                 E   K   F
LAV ELI                           EM                             K  E       F R

                410        420        430        440        450        460        470        480
LAV BRU  LHPDKWTVQP IVLPEKDSWT VNDIQKLVGK LNWASQIYPG IKVRQLCKLL RGTKALTEVI PLTEEAELEL AENREILKEP
ARV 2                M                            A     K
LAV MAL              Q  D E                             K         A    DIV    A
LAV ELI           S  K    E          N  ER

                490        500        510        520        530        540        550        560
LAV BRU  VHGVYYDPSK DLIAEIQKQG QGQWTYQIYQ EPFKNLKTGK YARTRGAHTN DVKQLTEAVQ KITTESIVIW GKTPKFKLPI
ARV 2      E          V                                  M                   VS          I
LAV MAL                                     QY           IKS                  AQ             R
LAV ELI                         H                        M            A     R S         R    R
```

FIG. 3C-2

```
                570        580        590        600        610        620        630        640

LAV BRU   QKETWETWWT EYWQATWIPE WEFVNTPPLV KLWYQLEKEP IVGAETFYVD GAASRETKLG KAGYVTNRGR QKVVTLTDTT
ARV 2           A  M                                                    N             D        SIA
LAV MAL         A                                 T                     N    K        D        S  E
LAV ELI            A                                   I                N             D        P

                650        660        670        680        690        700        710        720

LAV BRU   NQKTELQAIH LALQDSGLEV NIVTDSQYAL GIIQAQPDKS ESELVNQIIE QIIKKEKVYL AWVPAHKGIG GNEQVDKLVS
ARV 2                                                      S
LAV MAL                     S                                  I          Q D      S
LAV ELI            N

                730        740        750        760        770        780        790        800

LAV BRU   AGIRKVLFLD GIDKAQDEHE KYHSNWRAMA SDFNLPPVVA KEIVASCDKC QLKGEAMHGQ VDCSPGIWQL DCTHLEGKVI
ARV 2              N       E                                                                  I
LAV MAL   S                E                      I                                           I
LAV ELI   Q                E        N

                810        820        830        840        850        860        870        880

LAV BRU   LVAVHVASGY IEAEVIPAET GQETAYFLLK LAGRWPVKTI HTDNGSNFTS TTVKAACWWA GIKQEFGIPY NPQSQGVVES
ARV 2
LAV MAL   I                            I          VV              AA         N
LAV ELI                                           VV              AA
```

FIG. 3D-1

```
                890        900        910        920        930        940        950        960
LAV BRU  MNKELKKIIG QVRDQAEHLK TAVQMAVFIH NFKRKGGIGG YSAGERIVDI IATDIQTKEL QKQITKIQNF RVYYRDSRDP
ARV 2      N                                                                              KK
LAV MAL                E                                    I M                           N
LAV ELI                                      RR             I                 I

                970        980        990       1000       1010
LAV BRU  LWKGPAKLLW KGEGAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED
ARV 2                                                        G G
LAV MAL  I
LAV ELI  I                     K          V
```

FIG. 3D-2

```
ENV
                                        SP                                                     OMP

                       10         20         30          40         50         60         70         80
                                                    ↓
LAV BRU   MRVK---EKY QHLWRWGWKW GTMLLGILMI CSATEKLWVT VYYGVPVWKE ATTTLFCASD AKAYDTEVHN VWATHACVPT
ARV 2      K  GTRRN             --- - L    M                                   R
LAV MAL      REIQRN   NW        --- -M     M   T   IA D                             S E       I
LAV ELI     ARGIERNC  NW K      --- -I         T    ADN                             S E  A    I

                      90        100        110        120        130        140        150        160

LAV BRU   DPNPQEVVLV NVTENFNMWK NDMVEQMHED IISLWDQSLK PCVKLTPLCV SLKCTDL-CN ATNTNSSNTN SSSGEMMME-
ARV 2              C            N     Q                         T N    - K        ---  ----NWKE I
LAV MAL           IE E     G    N                               T N  NVN T  V GTNACS   RTNA LK  I
LAV ELI           IA E          N                               T N S E--L RN GTMG NV  TTEEKG----

                     170        180        190        200        210        220        230        240

LAV BRU   KGEIKNCSFN ISTSIRGKVQ KEYAFFYKLD IIPIDNDTTS -----YTLTS CNTSVITQAC PKVSFEPIPI HYCAPAGFAI
ARV 2                 T    D I     N L RN  VV     AS T TNYTN R IN   R                     T
LAV MAL   -  V        TPVGSD R  -   T  N   LVQ  DSDN   ----S R IN                  T D
LAV ELI   ---M       VT VLKD K  QV  L  R    V      SST -NSTN R IN      A

                     250        260        270        280        290        300        310        320

LAV BRU   LKCNKKTFNG TGPCTNVSTV QCTHGIRPVV STQLLLNGSL AEEEVVIRSA NFTDNAKTII VQLNQSVEIN CTRPNNNTRK
ARV 2                K                   I                      D     N          E  A
LAV MAL         D K    EI K          K                     IM    E  L   T N        ET T       G   R
LAV ELI        RD K                                        I     E  L N   N    AH  E  K T  A  YQ   Q
```

FIG. 3E–1

```
             330        340        350        360        370        380        390        400
                                  •                                                    •
BRU   SIRIQRGPGR AFVTIGK-IG NMRQAHCNIS RAKWNATLKQ IASKLREQFG NNKT-IIFKQ SSGGDPEIVT HSFNCGGEFF
2       Y --       W T RI   DI K         Q  N  E   VK            - V N           M      R
MAL   G HF --  Q  LY T I-V  DI R Y T N ETE DK  Q  V V  GSLL- -   K   NS         T       R
ELI   RTP -- L Q SLY TKS-RS IIG          Q SK  Q  V R  GTLL- -   I K  P         T

             410        420        430        440        450        460        470        480
       •                                   •                             •
BRU   YCNSTQLFNS TWFNSTWSTE CSNNTEGSDT ITLPCRIKQF INMWQEVGKA MYAPPISGQI RCSSNITGLL LTRDGGNN--
2        T     N   -----RLN RTEG K N    I       I                  C    S                T -V
MAL      TSK       Q NGARL- -  S  STGS          I      KT          A V  N L      I       NSSD
ELI      TSG       NI A NNI TES NSTNTN    Q     I  K VAGR-   I     ERN  L                I --

                                                                        TMP
             490        500        510        520        530        540        550        560
BRU   NNGSEIFRPG GGDMRDNWRS ELYKYKVVKI EPLGVAPTKA KRRVVQREKR AVGI-GALFL GFLGAAGSTM GARSMTLTVQ
2     T DT V                       I       I                     V  M                V L
MAL    SDN TL            I          R                  E      I L-  M                A L
ELI    STN T                        Q          R       E      I L-  M                  V
```

FIG. 3E–2

```
                   570        580        590        600        610        620        630        640
                                                                          •      •
LAV BRU     ARQLLSGIVQ QQNNLLRAIE AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICTTAVPW NASWSNKSLE
ARV 2                                                 W        R
LAV MAL                                               W        Q   R   M       H    F     S    R  D
LAV ELI         M                                                              H    N     S    R  N

                   650        660        670        680        690        700        710        720
LAV BRU     QIWNNMTWME WDREINNYTS LINSLIEESQ NQQEKNEQEL LELDKWASLW NWFNITNWLW YIKIFIMIVG GLVGLRIVFA
ARV 2       D  D     Q  E   D   N T YT L                              S
LAV MAL     D        Q  EK  S   G I YN       I      K                 S SK      R   IV     I    I
LAV ELI     E  Q        E   D   G   Y        T      K                 S  Q            I     I

                   730        740        750        760        770        780        790        800
                                                                                  •
LAV BRU     VLSIVNRVRQ GYSPLSFQTH LPTPRGP-DR PEGIEEEGGE RDRDRSIRLV NGSLALIWDD LRSLCLFSYH RLRDLLLIVT
ARV 2                           R   V    -    D               V    D F     E           R         AA
LAV MAL        L             L  L        P              QG G         FS         N                A
LAV ELI        L                L   A    -      T        G    V  L   FS                        I  AV

                   810        820        830        840        850        860        870
LAV BRU     RIVELLGRRG WEALKYWWNL LQYWSQELKN SAVSLLNATA IAVAEGTDRV IEVVQGACRA IRHIPRRIRQ GLERILL
ARV 2        T  I  K           S      I          W         T          A R Y    L  H          L
LAV MAL                      L        G        I    T        C       IG RFG    L          F  A
LAV ELI                 DI   L            R    S  FD I               II R     VLN            S
```

FIG. 3F–1

F

```
                      10          20          30          40          50          60          70          80
LAV BRU   MGGKWSKSSV VGWPTVRERM R----RAEPA ADGVGAASR- ----DLEKUG AITSSNTAAT NAACAWLEAQ EE-EEVGFPV
ARV 2             R M G   SAI       RAEP             V  -  ----                       D            -
LAV MAL             I      KI     I ----  TP  T ET      V QD AVSQ  D C    AA   SP  N   S --- PP   E
LAV ELI             I      AI     I ----  TM            V  -  ----                 S    D            SD

                      90         100         110         120         130         140         150         160
LAV BRU   TPQVPLRRHT YKAAVDLSHF LKEKGGLEGL IHSQRRQDIL DLWIYUTQGY FPDWQNYTPC PGVRYPLTFG WCYKLVPVEP
ARV 2     R                L  I              W     E                              I          F
LAV MAL   R              G F             D  VW  PK  E     V                         I F        F     HS
LAV ELI   R              E L                 W  KK  E     V N   I                   I           E    D

                     170         180         190         200         210
LAV BRU   DKVEEANKGE NTSLLHPVSL HGMDDPEREV LEWRFDSRLA FHHVARELHP EYFKNC
ARV 2     E        E   N     M     E A K   V    K      M          Y D
LAV MAL   EE       E   NC    I Q   E A     K K  S   LR R   Q      Y D
LAV ELI   QE    DTE   TN     ICQ   E   Q   K  N      E K   M     FY  -
```

FIG. 3F-2

FIG. 4A

A LAVbru vs.	GAG		POL		ENV					
					TOTAL		OMP		TMP	
HTLV-3 USA	512 0/0	0.8	1015 0/0	1.3	856 5/0	1.4	507 5/0	1.6	349 0/0	1.1
ARV-2 USA	502 12/2	3.4	1003 12/0	3.1	855 17/11	13.0	505 17/10	14.3	350 0/1	11.2
LAVeli ZAIRE	500 13/1	9.8	1002 13/0	5.5	853 22/14	20.7	504 22/14	25.3	349 0/0	13.8
LAVmal ZAIRE	505 14/7	12.0	1002 13/0	7.7	859 13/11	21.7	509 13/10	26.4	350 0/1	14.9

B LAVeli vs.										
LAVmal	505 1/6	10.8	1002 0/0	8.4	859 13/11	19.8	509 8/13	23.6	350 0/1	14.3

FIG. 4B

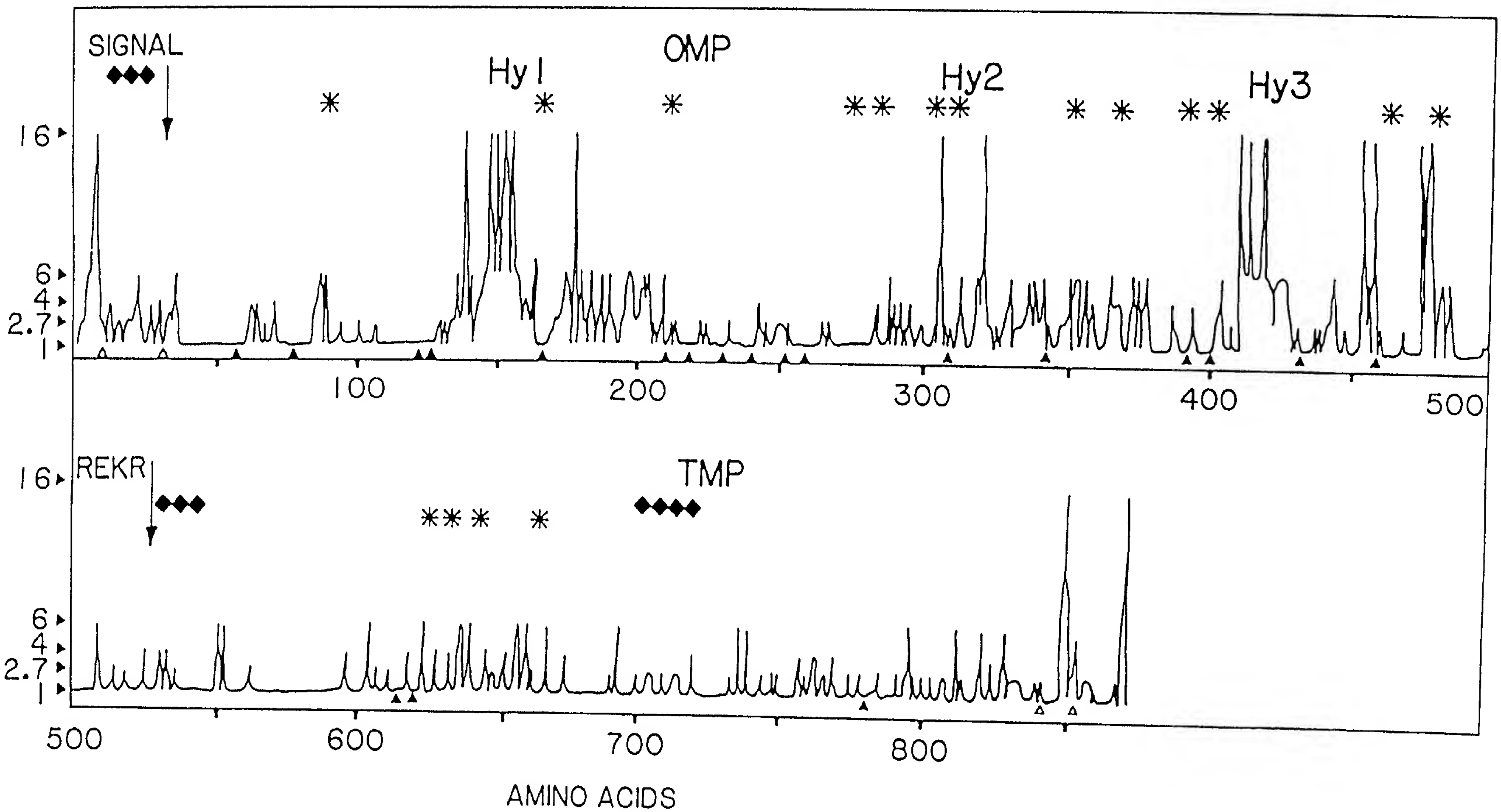
A LAVbru vs.

	orf F		central region					
			orf Q		orf R		orf S	
HTLV-3 USA	206 0/0	1.5	192 0/0	0		nd	80 0/0	2.5
ARV-2 USA	210 0/4	12.6	192 0/0	10.0	97 0/1	9.4	81 0/1	15.0
LAVeli ZAIRE	206 1/1	19.4	192 0/0	10.4	96 0/0	11.5	80 0/0	27.5
LAVmal ZAIRE	209 2/5	27.0	192 0/0	12.6	96 0/0	10.4	80 0/0	23.8

B LAVeli vs.

LAVmal	209 3/6	22.5	192 0/0	12.0	96 0/0	6.3	80 0/0	11.3

FIG. 5



SIGNAL
OMP
Hy1
Hy2
Hy3
REKR
TMP
16
6
4
2.7
1
100
200
300
400
500
600
700
800
AMINO ACIDS

GAG

a

				120											
	K	A	Q	Q	A	A	A							D	T
LAV.BRU	AAA	GCA	CAG	CAA	GCA	GCA	GCT	-	-	-	-	-		GAC	ACA
	K	A	Q	Q	A	A	A	A	A	-	-	-	-	G	T
ARV 2	AAG	GCA	CAG	CAA	GCA	GCA	GCT	GCA	GCT					GGC	ACA
	K	T	Q	Q	A	A	A	A	Q	Q	A	A	A	A	T
LAV.MAL	AAG	ACA	CAG	CAG	GCA	GCA	GCT	GCA	CAG	CAG	GCA	GCA	GCT	GCC	ACA
	X	A	Q	Q	A	A	A							D	T
LAV.ELI	AAG	GCA	CAG	CAA	GCA	GCA	GCT	-	-	-	-	-	-	GAC	ACA

FIG. 6A-1

```
.BRU
          460                                               470                                                  480
 N    F   L   Q   S   R   P   E   P   T   A   P   P     F   L   Q   S   R   P   E   P   T   A   P   P    E   E
AAT  TTT CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA   TTT CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA  GAA GAG

2
 N    F   L   Q   S   R   P   E   P   T   A   P   P                                                      E   E
AAT  TTT CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA   -   -   -   -   -   -   -   -   -   -   -   -    GAA GAG

.MAL
 N    F   L   Q   S   R   P   E   P   T   A   P   P                                                      A   E
AAT  TTC CTT CAG AGC AGA CCA GAG CCA ACA GCC CCA CCA   -   -   -   -   -   -   -   -   -   -   -   -    GCA GAG

.ELI
 N    F   L   Q   S   R   P   E   P   T   A   P   P                                                      A   E
AAC  TTT CTC CAA AGC AGA CCA GAG CCA ACA GCC CCA CCA   -   -   -   -   -   -   -   -   -   -   -   -    GCA GAG
```

FIG. 6A-2

```
c
                     20                                               30
LAV.BRU      R    M   .R                          R    A    E    P    A
            AGA  ATG  AGA   -    -    -    -   CGA  GCT  GAG  CCA  GCA

ARV 2        R    M    R    R    A    E    P    R    A    E    P    A
            AGA  ATG  AGA  CGA  GCT  GAG  CCA  CGA  GCT  GAG  CCA  GCA

LAV.MAL      R    I    R                          R    T    P    P    T
            AGA  ATA  AGA   -    -    -    -   CGA  ACT  CCC  CCA  ACA

LAV.ELI      R    I    R                          R    T    P    P    T
            AGA  ATA  AGA   -    -    -    -   AGA  ACT  AAT  CCA  GCA

d
                               40
LAV.BRU      V    G    A    A    S    R                                D
            GTG  GGA  GCA  GCA  TCT  CGA   -    -    -    -    -    - GAC

ARV 2        V    G    A    V    S    R                                D
            GTG  GGA  GCA  GTA  TCT  CGA   -    -    -    -    -    - GAC

LAV.MAL      V    G    A    V    S    R    D    A    V    S    Q    D
            GTA  GGA  GCA  GTA  TCT  CAA  GAT  GCA  GTA  TCT  CAA  GAT

LAV.ELI      V    G    A    V    S    R                                D
            GTA  GGA  GCA  GTA  TCT  CGA   -    -    -    -    -    - GAC
```

FIG. 6A–3

ENV

e

```
                                          20
           Q   H   L   W   R   W   G   W   K   W   G   T   M   L
LAV.BRU   CAG CAC TTG TGG ACA TGG GGC TGG AAA TGG GGC ACC ATG CTC

           Q   H   L   W   R   W   G                   T   L   L
ARV 2     CAG CAC TTG TGG AGA TGG GGC  -   -   -   -  ACC TTG CTC

           Q   N   W   W   R   W   G                   M   M   L
LAV.MAL   CAA AAC TGG TGG AGA TGG GGC  -   -   -   -  ATG ATG CTC

           Q   N   W   W   K   W   G                   T   M   L
LAV.ELI   CAA AAC TGG TGG AAA TCG GGC  -   -   -   -  ATC ATG CTC
```

FIG. 6B-1

f

```
LAV.BRU                            140                                     150
 L   K   C   T   D   L       G   N   A   T   N   T   N   S   S   N   T   N   S   S   S   G   E
TTA AAG TGC ACT GAT TTG  -  GGG AAT GCT ACT AAT ACC AAT AGT AGT AAT ACC AAT AGT AGT AGC GGG GAA
 M   M   M   E       K   G   E   I
ATG ATG ATG GAG  -  AAA GCA GAG ATA

ARV 2
 L   N   C   T   D   L       G   K   A   T   N   T   N   S   S                               M
TTA AAT TGC ACT GAT TTG  -  GGG AAG GCT ACT AAT ACC AAT AGT AGT  -   -   -   -   -   -   -  AAT
 W   K   E   E   I   K   G   E   I
TGG AAA GAA GAA ATA AAA GGA GAA ATA
```

```
LAV.MAL
 L   N   C   T  | N   V   N   G   T  | A   V   N   G   T  | N   A   G  | S   N   R   T   N   A   E
TTA AAC TGC ACT |AAT GTG AAT GGG ACT |GCT GTG AAT GGG ACT |AAT GCT GGG |AGT AAT AGG ACT AAT GCA GAA

 L   K   M   E   I       G   E   V
TTG AAA ATG GAA ATT  -  GGA GAA GTG

LAV.ELI
 L   N   C   S   D   E          | L   R   N   N   G   T  | M   G   N   N   V   T  | T   E   E   K
TTA AAC TGT AGT GAT GAA  -   -  |TTG AGG AAC AAT GGC ACT |ATG GGG AAC AAT GTC ACT |ACA GAG GAG AAA

 G                               M
GGA  -   -   -   -   -   -   -  ATG
```

FIG. 6B-2

```
g                                   200
               D   N   D   T   T   S                         Y   T   L
LAV.BRU       GAT AAT GAT ACT ACC AGC -   -   -   -   -     TAT ACG TTG

               D   N   A   S   T   T   T   N   Y   T   N   Y   R   L
ARV 2         GAT AAT GCT AGT ACT ACT ACC AAC TAT ACC AAC TAT AGG TTG

               D   D   S   D   N   S   S                     Y   R   L
LAV.MAL       GAT GAT AGT GAT AAT AGT AGT -   -   -   -     TAT AGG CTA

               D   N   D   S   S   T       N   S   T   N   Y   R   L
LAV.ELI       GAC AAT GAT AGT AGT ACC  -  AAT AGT ACC AAT TAT AGG TTA
```

FIG. 6B-3

```
h
LAV.BRU                           410                                          420
 C   N   S   T   Q   L   F   N   S   T   W   F   N   S   T   W   S   T   E   G   S   N   N   T   E   G
TGT AAT TCA ACA CAA CTG TTT AAT AGT ACT TGG TTT AAT AGT ACT TGG AGT ACT GAA GGG TCA AAT AAC ACT GAA GGA
        430
 S   D   T   I
AGT GAC ACA ATC

ARV 2
 C   N   T   T   Q   L   F   N   N   T   W                       R   L   N   H   T   K   G   T   K   G
TGT AAT ACA ACA CAA CTG TTT AAT AAT ACA TGG -   -   -   -   -   AGG TTA AAT CAC ACT GAA GGA ACT AAA GGA

 N   D   T   I
AAT GAC ACA ATC
```

```
LAV.MAL
 C   N   T   S   K   L   F   N   S   T   W   Q   N   N   G   A   R   L           S   N   S   T   E   S
TGT AAT ACA TCA AAA CTG TTT AAT AGT ACA TGG CAG AAT AAT GGT GCA AGA CTA  -   -  AGT AAT AGC ACA GAG TCA

 T   G   S   I
ACT GGT AGT ATC

LAV.ELI
 C   N   T   S   G   L   F   N   S   T   W   N   I   S   A   W   N   N   I   T   E   S   N   N   S   T
TGT AAT ACA TCA GGA CTG TTT AAT AGT ACA TGG AAT ATT AGT GCA TGG AAT AAT ATT ACA GAG TCA AAT AAT AGC ACA

 N   T   N   I
AAC ACA AAC ATC
```

FIG. 6B-4

LAV.ELI

→R
GGTCTCTCTGGTTAGACCAGATTTGAGCCTGGGAGCTCTCTGGCTAGCTAGGGAACCCAC
R ←→ U5
TGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGT
100
GTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGAGTGGAAAATCTCTAGCA
U5←
GTGGCGCCCGAACAGGGACCTGAAAGCGAAAGTAGAACCAGAGGAGCTCTCTCGACGCAG
200
GACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCAGCGACTGGTGAGTACGCT
→GAG. 300
MetGlyAlaArgAlaSerValLeuSer
AAAATTTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTATTAA

GlyGlyLysLeuAspLysTrpGluLysIleArgLeuArgProGlyGlyLysLysLysTyr
GCGGGGGAAAATTAGATAAATGGGAAAAAATTCGGTTACGGCCAGGAGGAAAGAAAAAAT
400
ArgLeuLysHisIleValTrpAlaSerArgGluLeuGluArgTyrAlaLeuAsnProGly
ATAGACTAAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATATGCACTTAATCCTG

LeuLeuGluThrSerGluGlyCysLysGlnIleIleGlyGlnLeuGlnProAlaIleGln
GCCTTTTAGAAACATCAGAAGGCTGTAAACAAATAATAGGGCAGCTACAACCAGCTATTC
500
ThrGlyThrGluGluLeuArgSerLeuTyrAsnThrValAlaThrLeuTyrCysValHis
AGACAGGAACAGAAGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTAC
600
LysGlyIleAspValLysAspThrLysGluAlaLeuGluLysMetGluGluGluGlnAsn
ATAAAGGAATAGATGTAAAAGACACCAAGGAAGCTTTAGAAAAGATGGAGGAAGAGCAAA

LysSerLysLysLysAlaGlnGlnAlaAlaAlaAspThrGlyAsnAsnSerGlnValSer
ACAAAAGTAAGAAAAAGGCACAGCAAGCAGCAGCTGACACAGGAAACAACAGCCAGGTCA
700
GlnAsnTyrProIleValGlnAsnLeuGlnGlyGlnMetValHisGlnAlaIleSerPro
GCCAAAATTATCCTATAGTGCAGAACCTACAGGGGCAAATGGTACATCAGGCCATATCAC

ArgThrLeuAsnAlaTrpValLysValIleGluGluLysAlaPheSerProGluValIle
CTAGAACTTTGAACGCATGGGTAAAAGTAATAGAAGAAAAGGCTTTCAGCCCAGAAGTAA
800
ProMetPheSerAlaLeuSerGluGlyAlaThrProGlnAspLeuAsnThrMetLeuAsn
TACCCATGTTTTCAGCATTATCAGAAGGAGCCACCCCACAAGATTTAAACACCATGCTAA
900
ThrValGlyGlyHisGlnAlaAlaMetGlnMetLeuLysGluThrIleAsnGluGluAla
ACACAGTGGGGGGACATCAAGCAGCCATGCAAATGCTAAAAGAGACCATCAATGAAGAAG

AlaGluTrpAspArgLeuHisProValHisAlaGlyProIleAlaProGlyGlnMetArg
CTGCAGAATGGGATAGGTTACATCCAGTGCATGCAGGGCCTATTGCACCAGGCCAGATGA
1000
GluProArgGlySerAspIleAlaGlyThrThrSerThrLeuGlnGluGlnIleAlaTrp
GAGAACCAAGGGGAAGTGATATAGCAGGAACTACTAGTACCCTTCAGGAACAAATAGCAT

MetThrSerAsnProProIleProValGlyGluIleTyrLysArgTrpIleIleValGly
GGATGACAAGTAACCCACCTATCCCAGTAGGAGAAATCTATAAAAGATGGATAATTGTGG
1100
LeuAsnLysIleValArgMetTyrSerProValSerIleLeuAspIleArgGlnGlyPro
GATTAAATAAAATAGTAAGAATGTATAGCCCTGTCAGCATTTTGGACATAAGACAGGGAC
1200

FIG. 7A

LysGluProPheArgAspTyrValAspArgPheTyrLysThrLeuArgAlaGluGlnAla
CAAAGGAACCTTTTAGAGACTATGTAGACCGGTTCTATAAAACTCTAAGAGCCGAGCAAG

SerGlnAspValLysAsnTrpMetThrGluThrLeuLeuValGlnAsnAlaAsnProAsp
CTTCACAGGATGTAAAAAATTGGATGACAGAAACCTTGTTGGTCCAAAATGCAAACCCAG
1300

CysLysThrIleLeuLysAlaLeuGlyProGlnAlaThrLeuGluGluMetMetThrAla
ATTGCAAGACTATCTTAAAAGCATTGGGACCACAGGCTACACTAGAAGAAATGATGACAG

CysGlnGlyValGlyGlyProSerHisLysAlaArgValLeuAlaGluAlaMetSerGln
CATGTCAGGGAGTGGGGGGCCCAGCCATAAAGCAAGAGTTCTGGCTGAGGCAATGAGCC
1400

AlaThrAsnSerValThrThrAlaMetMetGlnArgGlyAsnPheLysGlyProArgLys
AAGCAACAAATTCAGTTACTACAGCAATGATGCAGAGAGGCAATTTTAAGGGCCCAAGAA
1500

IleIleLysCysPheAsnCysGlyLysGluGlyHisIleAlaLysAsnCysArgAlaPro
AAATTATTAAGTGTTTCAATTGTGGCAAAGAAGGGCACATAGCAAAAAATTGCAGGGCCC

ArgLysLysGlyCysTrpArgCysGlyLysGluGlyHisGlnLeuLysAspCysThrGlu
CTAGGAAAAAGGGCTGTTGGAGATGTGGAAAGGAAGGACACCAACTAAAAGATTGCACTG
→POL 1600

PhePheArgGluAsnLeuAlaPheProGlnGlyLysAlaGlyGluLeu
ArgGlnAlaAsnPheLeuGlyArgIleTrpProSerHisLysGlyArgProGlyAsnPhe
AGAGACAGGCTAATTTTTTAGGGAGAATTTGGCCTTCCCACAAGGGAAGGCCGGGGAACT

SerProLysGlnThrArgAlaAsnSerProThrSerArgGluLeuArgValTrpGlyArg
LeuGlnSerArgProGluProThrAlaProProAlaGluSerPheGlyPheGlyGluGlu
TTCTCCAAAGCAGACCAGAGCCAACAGCCCCACCAGCAGAGAGCTTCGGGTTTGGGGAAG
1700

AspAsnProLeuSerLysThrGlyAlaGluArgGlnGlyThrValSerPheAsnPhePro
IleThrProSerGlnLysGlnGluGlnLysAspLysGluLeuTyrProLeuThrSerLeu
AGATAACCCCCTCTCAAAAACAGGAGCAGAAAGACAAGGAACTGTATCCTTTAACTTCCC
.GAG← 1800

GlnIleThrLeuTrpGlnArgProLeuValAlaIleLysIleGlyGlyGlnLeuLysGlu
LysSerLeuPheGlyAsnAspProLeuSerGln
TCAAATCACTCTTTGGCAACGACCCCTTGTCGCAATAAAAATAGGGGGACAGCTAAAGGA

AlaLeuLeuAspThrGlyAlaAspAspThrValLeuGluGluMetAsnLeuProGlyLys
AGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAATTTGCCAGGAAA
1900

TrpLysProLysMetIleGlyGlyIleGlyGlyPheIleLysValArgGlnTyrAspGln
ATGGAAACCAAAAATGATAGGGGGAATTGGAGGTTTTATCAAAGTAAGACAGTATGATCA

IleProIleGluIleCysGlyGlnLysAlaIleGlyThrValLeuValGlyProThrPro
AATACCCATAGAAATCTGTGGACAGAAAGCTATAGGTACAGTATTAGTAGGACCTACGCC
2000

ValAsnIleIleGlyArgAsnLeuLeuThrGlnIleGlyCysThrLeuAsnPheProIle
TGTCAACATAATCGGAAGAAATTTGTTGACCCAGATTGGCTGCACTTTAAATTTTCCAAT
2100

SerProIleGluThrValProValLysLeuLysProGlyMetAspGlyProLysValLys
TAGTCCTATTGAAACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAA

GlnTrpProLeuThrGluGluLysIleLysAlaLeuThrGluIleCysThrAspMetGlu
ACAATGGCCATTGACAGAAGAAAAAATAAAAGCATTAACAGAAATTTGTACAGATATGGA
2200

FIG. 7B

LysGluGlyLysIleSerArgIleGlyProGluAsnProTyrAsnThrProIlePheAla
AAAGGAAGGAAAAATTTCAAGAATTGGGCCTGAAAATCCATACAATACTCCAATATTTGC

IleLysLysLysAspSerThrLysTrpArgLysLeuValAspPheArgGluLeuAsnLys
CATAAAGAAAAAAGACAGTACCAAGTGGAGAAAATTAGTAGATTTCAGAGAACTTAATAA
2300

ArgThrGlnAspPheTrpGluValGlnLeuGlyIleProHisProAlaGlyLeuLysLys
GAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCGCATCCTGCAGGGCTGAAAAA
2400

LysLysSerValThrValLeuAspValGlyAspAlaTyrPheSerValProLeuAspGlu
GAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGA

AspPheArgLysTyrThrAlaPheThrIleSerSerIleAsnAsnGluThrProGlyIle
AGATTTTAGGAAATATACCGCCTTTACCATATCTAGTATAAACAATGAGACACCAGGGAT
2500

ArgTyrGlnTyrAsnValLeuProGlnGlyTrpLysGlySerProAlaIlePheGlnSer
TAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCGGCAATATTCCAAAG

SerMetThrLysIleLeuGluProPheArgLysGlnAsnProGluMetValIleTyrGln
TAGCATGACAAAAATCTTAGAGCCCTTTAGAAAACAAAATCCAGAAATGGTTATCTATCA
2600

TyrMetAspAspLeuTyrValGlySerAspLeuGluIleGlyGlnHisArgThrLysIle
ATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAGGGCAGCATAGGACAAAAAT
2700

GluLysLeuArgGluHisLeuLeuArgTrpGlyPheThrArgProAspLysLysHisGln
AGAGAAATTAAGAGAACATCTATTGAGGTGGGGATTTACCAGACCAGATAAAAAACATCA

LysGluProProPheLeuTrpMetGlyTyrGluLeuHisProAspLysTrpThrValGln
GAAAGAACCCCCATTTCTTTGGATGGGTTATGAACTCCATCCTGATAAATGGACAGTACA
2800

SerIleLysLeuProGluLysGluSerTrpThrValAsnAspIleGlnAsnLeuValGlu
GTCTATAAAACTGCCAGAAAAGGAGAGCTGGACTGTCAATGATATACAGAACTTAGTGGA

ArgLeuAsnTrpAlaSerGlnIleTyrProGlyIleLysValArgGlnLeuCysLysLeu
GAGATTAAACTGGGCAAGCCAGATTTATCCAGGAATTAAAGTAAGACAATTATGTAAACT
2900

LeuArgGlyThrLysAlaLeuThrGluValIleProLeuThrGluGluAlaGluLeuGlu
CCTTAGGGGAACCAAAGCACTAACAGAAGTAATACCACTAACAGAAGAAGCAGAATTAGA
3000

LeuAlaGluAsnArgGluIleLeuLysGluProValHisGlyValTyrTyrAspProSer
ACTGGCAGAAAACAGGGAAATTTTAAAAGAACCAGTACATGGAGTGTATTATGACCCATC

LysAspLeuIleAlaGluIleGlnLysGlnGlyHisGlyGlnTrpThrTyrGlnIleTyr
AAAAGACTTAATAGCAGAAATACAGAAACAAGGGCACGGCCAATGGACATACCAAATTTA
3100

GlnGluProPheLysAsnLeuLysThrGlyLysTyrAlaArgMetArgGlyAlaHisThr
TCAAGAACCATTTAAAAATCTGAAAACAGGAAAGTATGCAAGAATGAGGGGTGCCCACAC

AsnAspValLysGlnLeuAlaGluAlaValGlnArgIleSerThrGluSerIleValIle
TAATGATGTAAAGCAATTAGCAGAGGCAGTGCAAAGAATATCCACAGAAAGCATAGTGAT
3200

TrpGlyArgThrProLysPheArgLeuProIleGlnLysGluThrTrpGluThrTrpTrp
ATGGGGAAGGACTCCTAAATTTAGACTACCCATACAAAAGGAAACATGGGAAACATGGTG
3300

FIG. 7C

AlaGluTyrTrpGlnAlaThrTrpIleProGluTrpGluPheValAsnThrProProLeu
GGCAGAGTATTGGCAAGCCACTTGGATTCCTGAGTGGGAATTTGTCAATACCCCTCCTTT

ValLysLeuTrpTyrGlnLeuGluLysGluProIleIleGlyAlaGluThrPheTyrVal
AGTAAAATTATGGTACCAGTTAGAGAAGGAACCCATAATAGGAGCAGAAACTTTCTATGT
3400
AspGlyAlaAlaAsnArgGluThrLysLeuGlyLysAlaGlyTyrValThrAspArgGly
AGATGGGGCAGCTAATAGAGAGACTAAATTAGGAAAAGCAGGATATGTTACTGACAGAGG

ArgGlnLysValValProLeuThrAspThrThrAsnGlnLysThrGluLeuGlnAlaIle
AAGACAGAAAGTTGTCCCTTTGACTGACACGACAAATCAGAAGACTGAGTTACAAGCAAT
3500
AsnLeuAlaLeuGlnAspSerGlyLeuGluValAsnIleValThrAspSerGlnTyrAla
TAATCTAGCCTTGCAGGATTCGGGATTAGAAGTAAACATAGTAACAGATTCACAATATGC
3600
LeuGlyIleIleGlnAlaGlnProAspLysSerGluSerGluLeuValAsnGlnIleIle
ATTAGGAATCATTCAAGCACAACCAGATAAGAGTGAATCAGAGTTAGTCAATCAAATAAT

GluGlnLeuIleLysLysGluLysValTyrLeuAlaTrpValProAlaHisLysGlyIle
AGAGCAGTTAATAAAAAAGGAAAAGGTTTACCTGGCATGGGTACCAGCACACAAAGGAAT
3700
GlyGlyAsnGluGlnValAspLysLeuValSerGlnGlyIleArgLysValLeuPheLeu
TGGAGGAAATGAACAAGTAGATAAATTAGTCAGTCAAGGAATCAGGAAAGTACTATTTTT

AspGlyIleAspLysAlaGlnGluGluHisGluLysTyrHisAsnAsnTrpArgAlaMet
GGATGGAATAGATAAGGCTCAAGAAGAACATGAGAAATATCACAACAATTGGAGAGCAAT
3800
AlaSerAspPheAsnLeuProProValValAlaLysGluIleValAlaSerCysAspLys
GGCTAGTGATTTTAACCTACCACCCGTGGTAGCAAAAGAAATAGTAGCTAGCTGTGATAA
3900
CysGlnLeuLysGlyGluAlaMetHisGlyGlnValAspCysSerProGlyIleTrpGln
ATGTCAGCTAAAAGGAGAAGCCATGCATGGACAAGTAGACTGTAGTCCAGGAATATGGCA

LeuAspCysThrHisLeuGluGlyLysValIleLeuValAlaValHisValAlaSerGly
ATTAGATTGTACACACTTAGAAGGAAAAGTTATCCTGGTAGCAGTTCATGTAGCCAGTGG
4000
TyrIleGluAlaGluValIleProAlaGluThrGlyGlnGluThrAlaTyrPheLeuLeu
CTATATAGAAGCAGAAGTTATTCCAGCAGAAACAGGGCAGGAAACAGCATATTTTCTTTT

LysLeuAlaGlyArgTrpProValLysValValHisThrAspAsnGlySerAsnPheThr
AAAATTAGCAGGAAGATGGCCAGTAAAAGTAGTACATACAGACAATGGCAGCAATTTCAC
4100
SerAlaAlaValLysAlaAlaCysTrpTrpAlaGlyIleLysGlnGluPheGlyIlePro
CAGTGCTGCAGTTAAGGCCGCCTGTTGGTGGGCAGGTATCAAACAGGAATTTGGAATTCC
4200
TyrAsnProGlnSerGlnGlyValValGluSerMetAsnLysGluLeuLysLysIleIle
CTACAATCCCCAAAGTCAAGGAGTAGTAGAATCTATGAATAAAGAATTAAAGAAAATTAT

GlyGlnValArgAspGlnAlaGluHisLeuLysThrAlaValGlnMetAlaValPheIle
AGGACAGGTAAGAGATCAAGCTGAACATCTTAAGACAGCAGTACAAATGGCAGTATTCAT
4300
HisAsnPheLysArgArgArgGlyIleGlyGlyTyrSerAlaGlyGluArgIleIleAsp
CCACAATTTTAAAAGAAGAAGGGGGATTGGGGGATACAGTGCAGGGGAAAGAATAATAGA

FIG. 7D

```
IleIleAlaThrAspIleGlnThrLysGluLeuGlnLysGlnIleIleLysIleGlnAsn
CATAATAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTATAAAAATTCAAAA
                  4400
PheArgValTyrTyrArgAspSerArgAspProIleTrpLysGlyProAlaLysLeuLeu
TTTTCGGGTTTATTACAGAGACAGCAGAGATCCAATTTGGAAAGGACCAGCAAAGCTCCT
                                                        4500
TrpLysGlyGluGlyAlaValValIleGlnAspLysSerAspIleLysValValProArg
CTGGAAAGGTGAAGGGGCAGTAGTAATACAAGACAAGAGTGACATAAAGGTAGTACCAAG
                         →Q
ArgLysValLysIleIleArgAspTyrGlyLysGlnMetAlaGlyAspAspCysValAla
                         MetGluAsnArgTrpGlnValMetIleValTrpGln
AAGAAAAGTAAAGATTATTAGGGATTATGGAAAACAGATGGCAGGTGATGATTGTGTGGC
                 POL←                     4600
SerArgGlnAspGluAsp
 ValAspArgMetArgIleLysThrTrpLysSerLeuValLysHisHisMetTyrValSer
AAGTAGACAGGATGAGGATTAAAACATGGAAAAGTTTAGTAAAACACCATATGTATGTTT

 LysLysAlaAsnArgTrpPheTyrArgHisHisTyrGluSerProHisProLysIleSer
CAAAGAAAGCTAACAGATGGTTTTATAGACATCACTATGAAAGCCCCCACCCAAAAATAA
                  4700
 SerGluValHisIleProLeuGlyGluAlaArgLeuValIleLysThrTyrTrpGlyLeu
GTTCAGAAGTACACATCCCACTAGGAGAAGCTAGACTGGTAATAAAAACATATTGGGGTC
                                                        4800
 HisThrGlyGluArgGluTrpHisLeuGlyGlnGlyValSerIleGluTrpArgLysArg
TGCATACAGGAGAAAGAGAATGGCATCTGGGTCAGGGAGTCTCCATAGAATGGAGGAAAA

 ArgTyrSerThrGlnValAspProGlyLeuAlaAspGlnLeuIleHisMetTyrTyrPhe
GGAGATATAGCACACAAGTAGACCCTGGCCTGGCAGACCAACTAATTCATATGTATTATT
                                          4900
 AspCysPheSerGluSerAlaIleArgLysAlaIleLeuGlyAspIleValSerProArg
TTGATTGTTTTTCAGAATCTGCTATAAGAAAAGCCATATTAGGAGATATAGTTAGTCCTA

 CysGluTyrGlnAlaGlyHisAsnLysValGlySerLeuGlnTyrLeuAlaLeuThrAla
GGTGTGAGTATCAAGCAGGACATAACAAGGTAGGATCCCTACAGTATTTGGCACTAACAG
                  5000
 LeuIleAlaProLysGlnIleLysProProLeuProSerValArgLysLeuThrGluAsp
CATTAATAGCACCAAAACAGATAAAGCCACCTTTGCCTAGTGTTAGGAAGCTAACAGAAG
   →R                                                   5100
   MetGluGlnAlaProAlaAspGlnGlyProGlnArgGluProTyrAsnGluTrpAla
 ArgTrpAsnLysProGlnGlnThrArgGlyHisArgGlySerHisThrMetAsnGlyHis
ATAGATGGAACAAGCCCCAGCAGACCAGGGGCCACAGAGGGAGCCATACAATGAATGGGC
Q←
 LeuGluLeuLeuGluGluLeuLysSerGluAlaValArgHisPheProArgIleTrpLeu
ATTAGAGCTTTTAGAGGAGCTTAAGAGTGAAGCTGTTAGACATTTTCCTAGGATATGGCT
                                          5200
 HisSerLeuGlyGlnHisIleTyrGluThrTyrGlyAspThrTrpValGlyValGluAla
CCATAGCTTAGGACAACATATTTATGAAACTTATGGGGATACCTGGGTAGGAGTTGAAGC

 IleIleArgIleLeuGlnGlnLeuLeuPheIleHisPheArgIleGlyCysGlnHisSer
TATAATAAGAATACTGCAACAATTACTGTTTATTCATTTCAGAATTGGGTGTCAACATAG
                  5300                     →S          R←
 ArgIleGlyIleIleArgGlnArgArgAlaArgAsnGlySerSerArgSer
                                  MetAspProValAspProAsnLeuGlu
CAGAATAGGCATTATTCGACAGAGAAGAGCAAGAAATGGATCCAGTAGATCCTAACCTAG
                                                        5400
```

FIG. 7E

```
ProTrpAsnHisProGlySerGlnProArgThrProCysAsnLysCysHisCysLysLys
AGCCCTGGAACCATCCAGGAAGTCAGCCTAGGACTCCTTGTAACAAGTGTCATTGTAAAA

CysCysTyrHisCysProValCysPheLeuAsnLysGlyLeuGlyIleSerTyrGlyArg
AGTGTTGCTATCATTGCCCAGTTTGCTTCTTAAACAAAGGCTTAGGCATCTCCTATGGCA
                                        5500
LysLysArgArgGlnArgArgGlyProProGlnGlyGlyGlnAlaHisGlnValProIle
GGAAGAAGCGGAGACAGCGACGAGGACCTCCTCAAGGCGGTCAGGCTCATCAAGTTCCTA
      S
ProLysGln
TACCAAAGCAGTAAGTAGTACATGTAATGCAACCTTTAGGGATAATAGCAATAGCAGCAT
                    5600
TAGTAGTAGCAATAATACTAGCAATAGTTGTGTGGACCATAGTATTCATAGAATATAGAA
                                                        5700
GGATAAAAAAGCAAAGGAGAATAGACTGTTTACTTGATAGAATAACAGAAAGAGCAGAAG
          ENV
          MetArgAlaArgGlyIleGluArgAsnCysGlnAsnTrpTrpLysTrpGly
ACAGTGGCAATGAGAGCGAGGGGGATAGAGAGAAATTGTCAAAACTGGTGGAAATGGGGC
                                        5800
IleMetLeuLeuGlyIleLeuMetThrCysSerAlaAlaAspAsnLeuTrpValThrVal
ATCATGCTCCTTGGGATATTGATGACCTGTAGTGCTGCAGACAATCTGTGGGTCACAGTT

TyrTyrGlyValProValTrpLysGluAlaThrThrThrLeuPheCysAlaSerAspAla
TATTATGGGGTGCCTGTATGGAAGGAAGCAACCACCACTCTATTTTGTGCATCAGATGCT
                    5900
LysSerTyrGluThrGluAlaHisAsnIleTrpAlaThrHisAlaCysValProThrAsp
AAATCATATGAAACAGAGGCACATAATATCTGGGCCACACATGCCTGTGTACCCACGGAC
                                                        6000
ProAsnProGlnGluIleAlaLeuGluAsnValThrGluAsnPheAsnMetTrpLysAsn
CCCAACCCACAAGAAATAGCACTGGAAAATGTGACAGAAAACTTTAACATGTGGAAAAAT

AsnMetValGluGlnMetHisGluAspIleIleSerLeuTrpAspGlnSerLeuLysPro
AACATGGTGGAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAACCA
                                        6100
CysValLysLeuThrProLeuCysValThrLeuAsnCysSerAspGluLeuArgAsnAsn
TGTGTAAAATTAACCCCACTCTGTGTCACTTTAAACTGTAGTGATGAATTGAGGAACAAT

GlyThrMetGlyAsnAsnValThrThrGluGluLysGlyMetLysAsnCysSerPheAsn
GGCACTATGGGGAACAATGTCACTACAGAGGAGAAAGGAATGAAAAACTGCTCTTTCAAT
                    6200
ValThrThrValLeuLysAspLysLysGlnGlnValTyrAlaLeuPheTyrArgLeuAsp
GTAACCACAGTACTAAAAGATAAGAAGCAGCAAGTATATGCACTTTTTTATAGACTTGAT
                                                        6300
IleValProIleAspAsnAspSerSerThrAsnSerThrAsnTyrArgLeuIleAsnCys
ATAGTACCAATAGACAATGATAGTAGTACCAATAGTACCAATTATAGGTTAATAAATTGT

AsnThrSerAlaIleThrGlnAlaCysProLysValSerPheGluProIleProIleHis
AATACCTCAGCCATTACACAGGCTTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACAT
                                        6400
TyrCysAlaProAlaGlyPheAlaIleLeuLysCysArgAspLysLysPheAsnGlyThr
TATTGTGCCCCAGCTGGTTTTGCGATTCTAAAGTGTAGAGATAAGAAGTTCAATGGAACA

GlyProCysThrAsnValSerThrValGlnCysThrHisGlyIleArgProValValSer
GGCCCATGCACAAATGTCAGCACAGTACAATGTACACATGGAATTAGGCCAGTGGTGTCA
                    6500
```

FIG. 7F

```
ThrGlnLeuLeuLeuAsnGlySerLeuAlaGluGluGluValIleIleArgSerGluAsn
ACTCAACTGCTGTTGAATGGCAGTCTAGCAGAAGAAGAGGTCATAATTAGATCCGAAAAT
                                                        6600
LeuThrAsnAsnAlaLysAsnIleIleAlaHisLeuAsnGluSerValLysIleThrCys
CTCACAAACAATGCTAAAAACATAATAGCACATCTTAATGAATCTGTAAAAATTACCTGT

AlaArgProTyrGlnAsnThrArgGlnArgThrProIleGlyLeuGlyGlnSerLeuTyr
GCAAGGCCCTATCAAAATACAAGACAAAGAACACCTATAGGACTAGGGCAATCACTCTAT
                                      6700
ThrThrArgSerArgSerIleIleGlyGlnAlaHisCysAsnIleSerArgAlaGlnTrp
ACTACAAGATCAAGATCAATAATAGGACAAGCACATTGTAATATTAGTAGAGCACAATGG

SerLysThrLeuGlnGlnValAlaArgLysLeuGlyThrLeuLeuAsnLysThrIleIle
AGTAAAACTTTACAACAAGTAGCTAGAAAATTAGGAACCCTTCTTAACAAAACAATAATA
                  6800
LysPheLysProSerSerGlyGlyAspProGluIleThrThrHisSerPheAsnCysGly
AAGTTTAAACCATCCTCAGGAGGGGACCCAGAAATTACAACACACAGTTTTAATTGTGGA
                                                        6900
GlyGluPhePheTyrCysAsnThrSerGlyLeuPheAsnSerThrTrpAsnIleSerAla
GGGGAATTCTTCTACTGTAATACATCAGGACTGTTTAATAGTACATGGAATATTAGTGCA

TrpAsnAsnIleThrGluSerAsnAsnSerThrAsnThrAsnIleThrLeuGlnCysArg
TGGAATAATATTACAGAGTCAAATAATAGCACAAACACAAACATCACACTCCAATGCAGA
                                      7000
IleLysGlnIleIleLysMetValAlaGlyArgLysAlaIleTyrAlaProProIleGlu
ATAAAACAAATTATAAAGATGGTGGCAGGCAGGAAAGCAATATATGCCCCTCCTATCGAA

ArgAsnIleLeuCysSerSerAsnIleThrGlyLeuLeuLeuThrArgAspGlyGlyIle
AGAAACATTCTATGTTCATCAAATATTACAGGGCTACTATTGACAAGAGATGGTGGTATA
                  7100
AsnAsnSerThrAsnGluThrPheArgProGlyGlyGlyAspMetArgAspAsnTrpArg
AATAATAGTACTAACGAGACCTTTAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA
                                                        7200
SerGluLeuTyrLysTyrLysValValGlnIleGluProLeuGlyValAlaProThrArg
AGTGAATTATATAAATATAAGGTAGTACAAATTGAACCACTAGGAGTAGCACCCACCAGG

AlaLysArgArgValValGluArgGluLysArgAlaIleGlyLeuGlyAlaMetPheLeu
GCAAAGAGAAGAGTGGTGGAAAGAGAAAAAAGAGCAATAGGATTAGGAGCTATGTTCCTT
                                      7300
GlyPheLeuGlyAlaAlaGlySerThrMetGlyAlaArgSerValThrLeuThrValGln
GGGTTCTTGGGAGCAGCAGGAAGCACGATGGGCGCACGGTCAGTGACGCTGACGGTACAG

AlaArgGlnLeuMetSerGlyIleValGlnGlnGlnAsnAsnLeuLeuArgAlaIleGlu
GCCAGACAATTAATGTCTGGTATAGTGCAACAGCAAAACAATTTGCTGAGGGCTATAGAG
                  7400
AlaGlnGlnHisLeuLeuGlnLeuThrValTrpGlyIleLysGlnLeuGlnAlaArgIle
GCGCAACAGCATCTGTTGCAACTCACGGTCTGGGGCATTAAACAGCTCCAGGCAAGAATC
                                                        7500
LeuAlaValGluArgTyrLeuLysAspGlnGlnLeuLeuGlyIleTrpGlyCysSerGly
CTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTAGGAATTTGGGGTTGCTCTGGA
```

FIG. 7G

LysHisIleCysThrThrAsnValProTrpAsnSerSerTrpSerAsnArgSerLeuAsn
AAACACATTTGCACCACTAATGTGCCCTGGAACTCTAGTTGGAGTAATAGATCTCTAAAT
7600

GluIleTrpGlnAsnMetThrTrpMetGluTrpGluArgGluIleAspAsnTyrThrGly
GAGATTTGGCAGAACATGACCTGGATGGAGTGGGAAAGAGAAATTGACAATTACACAGGC

LeuIleTyrSerLeuIleGluGluSerGlnThrGlnGlnGluLysAsnGluLysGluLeu
TTAATATATAGCTTAATTGAGGAATCGCAGACCCAGCAAGAAAAGAATGAAAAAGAATTG
7700

LeuGluLeuAspLysTrpAlaSerLeuTrpAsnTrpPheSerIleThrGlnTrpLeuTrp
TTGGAATTGGACAAGTGGGCAAGTTTGTGGAATTGGTTTAGCATAACACAATGGCTGTGG
7800

TyrIleLysIlePheIleMetIleIleGlyGlyLeuIleGlyLeuArgIleValPheAla
TATATAAAAATATTCATAATGATAATAGGAGGCTTGATAGGTTTAAGAATAGTTTTTGCT

ValLeuSerLeuValAsnArgValArgGlnGlyTyrSerProLeuSerPheGlnThrLeu
GTGCTTTCTTTAGTAAATAGAGTTAGGCAGGGATACTCACCTCTGTCGTTTCAGACCCTC
7900

LeuProAlaProArgGlyProAspArgProGluGlyThrGluGluGluGlyGlyGluArg
CTCCCAGCCCCGAGGGGACCCGACAGGCCCGAAGGAACAGAAGAAGAAGGTGGAGAGCGA

GlyArgAspArgSerValArgLeuLeuAsnGlyPheSerAlaLeuIleTrpAspAspLeu
GGCAGAGACAGATCCGTGAGATTGCTGAACGGATTCTCGGCACTTATCTGGGACGACCTG
8000

ArgSerLeuCysLeuPheSerTyrHisArgLeuArgAspLeuIleLeuIleAlaValArg
CGGAGCCTGTGCCTCTTCAGCTACCACCGCTTGAGAGACTTAATCTTAATTGCAGTGAGG
8100

IleValGluLeuLeuGlyArgArgGlyTrpAspIleLeuLysTyrLeuTrpAsnLeuLeu
ATTGTAGAACTTCTGGGACGCAGGGGGTGGGACATCCTCAAATATCTGTGGAATCTCCTA

GlnTyrTrpSerGlnGluLeuArgAsnSerAlaSerSerLeuPheAspAlaIleAlaIle
CAGTATTGGAGTCAGGAACTGAGGAACAGTGCTAGTAGCTTGTTTGATGCCATAGCAATA
8200

AlaValAlaGluGlyThrAspArgValIleGluIleIleGlnArgAlaCysArgAlaVal
GCAGTAGCTGAGGGGACAGATAGAGTTATAGAAATAATACAAAGAGCTTGCAGAGCTGTT

W←

LeuAsnIleProArgArgIleArgGlnGlyLeuGluArgSerLeuLeu

→F

MetGlyGly
CTTAACATACCCAGAAGAATAAGACAGGGCTTAGAAAGGTCTTTACTTTAAAATGGGTGG
8300

LysTrpSerLysSerSerIleValGlyTrpProAlaIleArgGluArgIleArgArgThr
CAAATGGTCAAAAAGTAGTATAGTGGGATGGCCTGCTATAAGGGAAAGAATAAGAAGAAC
8400

AsnProAlaAlaAspGlyValGlyAlaValSerArgAspLeuGluLysHisGlyAlaIle
TAATCCAGCAGCAGATGGGGTAGGAGCAGTATCTCGAGACCTGGAAAAACATGGGGCAAT

ThrSerSerAsnThrAlaSerThrAsnAlaAspCysAlaTrpLeuGluAlaGlnGluGlu
CACAAGTAGCAATACAGCAAGTACTAATGCTGACTGTGCCTGGCTAGAAGCACAAGAAGA
8500

SerAspGluValGlyPheProValArgProGlnValProLeuArgProMetThrTyrLys
GAGCGACGAGGTGGGCTTTCCAGTCAGACCCCAGGTACCTTTAAGACCAATGACTTACAA

→U3

GluAlaLeuAspLeuSerHisPheLeuLysGluLysGlyGlyLeuGluGlyLeuIleTrp
AGAAGCTCTAGATCTCAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTTG
8600

FIG. 7H

SerLysLysArgGlnGluIleLeuAspLeuTrpValTyrAsnThrGlnGlyIlePhePro
GTCCAAAAAGAGACAAGAGATCCTTGATCTTTGGGTCTACAACACACAAGGCATCTTCCC
8700
AspTrpGlnAsnTyrThrProGlyProGlyIleArgTyrProLeuThrPheGlyTrpCys
TGATTGGCAAAACTACACACCAGGGCCAGGGATCAGATATCCACTAACCTTTGGATGGTG

TyrGluLeuValProValAspProGlnGluValGluGluAspThrGluGlyGluThrAsn
CTACGAGCTAGTACCAGTTGATCCACAGGAGGTAGAAGAAGACACTGAAGGAGAGACCAA
8800
SerLeuLeuHisProIleCysGlnHisGlyMetGluAspProGluArgGlnValLeuLys
CAGCTTGTTACACCCTATATGCCAGCATGGAATGGAGGACCCGGAGAGACAAGTGTTAAA

TrpArgPheAsnSerArgLeuAlaPheGluHisLysAlaArgGluMetHisProGluPhe
ATGGAGATTTAACAGCAGACTAGCATTTGAGCACAAGGCCCGAGAGATGCATCCGGAGTT
8900
TyrLysAsn
CTACAAAAACTGATGACACCGAGCTTTCTACAAGGGACTTTCCGCTGGGGACTTTCCAGG
9000
GAGGCGTGGACTGGGCGGGACTGGGGAGTGGCTAACCCTCAGATGCTGCATATAAGCAGC
U3 ← → R
TGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATTTGAGCCTGGGAGCTCTCTG
9100 B ←
GCTAGCTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAA

FIG. 7I